ADVANCED TRANSLATION OF SCIENTIFIC TERMINOLOGY USING COMPUTATIONAL LINGUISTICS: THEORY, PRACTICE, AND LIMITATIONS

G‘afurova Nazokat Bakhriddin’s daughter
student of Tashkent State of transport university

Annotation: This academic paper examines the translation of scientific terminology through computational linguistic tools. Emphasizing theoretical insights and practical comparisons, it evaluates the capabilities and limitations of machine translation (MT), computer-assisted translation (CAT) tools, and terminological databases in handling domain-specific vocabulary. Drawing on translation theory (Newmark, 1988; Vinay & Darbelnet, 1958; Vermeer, 1989) and empirical examples, it argues that while technology improves efficiency, human expertise remains essential for semantic and contextual accuracy.

Keywords: Computational linguistics, machine translation, CAT tools, scientific terminology, semantic accuracy, translation theory.

Annotatsiya: Ushbu maqolada ilmiy atamalarni tarjima qilishda kompyuter lingvistikasi vositalaridan foydalanishning nazariy va amaliy jihatlari o‘rganiladi. Mashinaviy tarjima (MT), tarjima xotirasi tizimlari (CAT), hamda terminologik ma’lumotlar bazalarining imkoniyatlari va chegaralari tahlil qilinadi. Newmark, Vinay va Darbelnet, hamda Vermeer nazariyalari asosida ilmiy atamalarni kontekstda to‘g‘ri va adekvat tarjima qilish uchun zamonaviy yondashuvlar taklif qilinadi. Amaliy misollar orqali tarjima vositalari o‘rtasidagi farqlar ko‘rsatilib, yuqori sifatli tarjima uchun gibrid yondashuvlar tavsiya etiladi.

Kalit so‘zlar: kompyuter lingvistikasi, ilmiy atamalar, mashinaviy tarjima, tarjima nazariyasi, kontekstual muammolar, tarjima xotirasi.

Аннотация: В данной статье рассматриваются теоретические и практические аспекты использования средств компьютерной лингвистики при переводе научной терминологии. Анализируются возможности и ограничения машинного перевода (MT), инструментов компьютерной поддержки перевода (CAT) и терминологических баз данных. На основе теорий Ньюмарка, Вине и Дарбельне, а также Вермеера предлагаются современные подходы к адекватному переводу терминов в контексте. Сравнительный анализ демонстрирует различия между инструментами перевода, предлагая гибридные стратегии для повышения качества перевода.

Ключевые слова: компьютерная лингвистика, научная терминология, машинный перевод, теория перевода, контекстуальные проблемы, память перевода.

The globalization of science necessitates accurate and context-sensitive translation of scientific terminology across languages. The complexity of domain-specific vocabulary often presents challenges for both human and machine translators. Computational linguistics has introduced powerful tools, but the quality of output remains uneven, especially in fields requiring precision such as medicine, law, and engineering (*Newmark, 1988*). This paper investigates the current capabilities of such tools and offers a theoretically grounded approach to their integration into professional workflows.

Translation theory has evolved to address not only linguistic transfer but also contextual, pragmatic, and cultural dimensions. *Newmark (1988)* differentiates between semantic translation (focused on meaning) and communicative translation (focused on effect). *Vinay and Darbelnet (1958)* identify translation techniques such as borrowing, calque, and modulation, essential for adapting technical terminology. *Vermeer's Skopos theory (1989)* emphasizes the purpose of translation as central to determining strategy, particularly in scientific communication where accuracy outweighs stylistic preferences.

Classification of Computational Linguistic Tools

Scientific terminology translation increasingly depends on three categories of tools:

- **Machine Translation (MT)**: e.g., Google Translate, DeepL
- **Computer-Assisted Translation (CAT)**: e.g., SDL Trados, MemoQ, Smartcat
- **Terminological Databases**: e.g., IATE, Termium, UNTerm

MT systems like Google Translate provide rapid outputs, but as *Hutchins (2005)* notes, they lack contextual nuance. CAT tools enhance consistency using translation memory (TM) and terminology management. Databases such as IATE provide validated equivalents to ensure terminological uniformity (*Bowker & Fisher, 2010*).

Main Challenges in Terminology Translation

Despite advancements, several challenges persist:

- **Polysemy**: Terms such as *"network"* vary in meaning across disciplines (e.g., IT vs. biology).
- **Cultural absence**: Concepts like *"peer review"* may lack direct equivalents in certain cultures.
- **Semantic drift**: MT systems may dilute precise meaning (e.g., *"antibiotic resistance"* translated as *"qarshilik"* instead of *"rezistentlik"*).
- **Inconsistency**: Different tools yield varied outputs for identical terms (*Koehn, 2020*).

Practical Comparison of Tools

Term	Google translate	DeepL	SDL Trados
Cloud computing	Bulutli hisoblash	Bulut asosida hisoblash	Bulutli texnologiyalar asosidagi hisoblash
Antibiotic resistance	Antibiotik qarshiligi	Antibiotikga chidamlilik	Antibiotikga rezistentlik
Emission control system	Emissiya nazorati tizimi	Emissiya boshqaruv tizimi	Chiqarilayotgan gazlarni boshqarish tizimi

As shown in the table, SDL Trados—especially when supported by subject-specific TMs—offers more nuanced and contextually accurate translations. The phrase *"Emission control system"* is particularly well-rendered in Trados, as it reflects the terminology used in environmental and automotive documentation (*Bowker & Fisher, 2010*).

Analysis of Comparative Results

As demonstrated in the table, automated translation systems—particularly Google Translate and DeepL—offer rapid and generally intelligible outputs; however, they fall short in domains that demand terminological precision and contextual accuracy. This observation aligns with *Newmark's (1988)* assertion that translation is not merely linguistic substitution but involves complex semantic and pragmatic transfer. For instance, the term **"antibiotic resistance"**, when rendered as *"qarshilik"* (resistance), may suffice in general discourse but fails to convey the technical nuance intended by *"rezistentlik"* in Uzbek medical terminology.

Moreover, while DeepL occasionally offers context-sensitive suggestions superior to Google Translate, it often exhibits a tendency toward **stylistic simplification** or **semantic generalization**. A good example is the translation of **"cloud computing"** as *"bulut asosida hisoblash"* (computing based on clouds), which is grammatically correct but not widely used or recognized in Uzbek technical discourse. In contrast, SDL Trados translates the same term as *"bulutli texnologiyalar asosidagi hisoblash"*, aligning with established professional usage. This reflects *Vinay and Darbelnet's (1958)* concepts of modulation and calque when appropriately applied through domain-aware memory banks.

Notably, SDL Trados, supported by specialized translation memory (TM) and integrated terminology management systems, produces highly **domain-specific and semantically appropriate translations**. For example, the phrase **"emission control system"** is rendered in Trados as *"chiqarilayotgan gazlarni boshqarish tizimi"*, which

fully captures the environmental and automotive regulatory context (*Bowker & Fisher, 2010*).

Another critical issue pertains to **cultural connotation** and the translation of **realia**—terms deeply embedded in the cultural and institutional fabric of the source language. Phrases such as *"peer review"* or *"tenure track"* often lack direct equivalents in Uzbek and must be handled via **descriptive translation** or **conceptual reformulation**. According to *Vermeer's Skopos Theory (1989)*, such cases demand a purpose-driven strategy where the translator adapts the form to match the communicative function in the target culture.

Furthermore, modern neural MT systems, including DeepL, employ **corpus-based probabilistic algorithms** that may generate fluent yet **falsely confident translations**—outputs that are grammatically accurate but semantically misleading. This phenomenon reinforces *Koehn's (2020)* argument that human post-editing remains indispensable, especially in specialized scientific and technical communication.

In summary, while computational tools are indispensable for scaling translation productivity, their effectiveness is maximized only when used in conjunction with human expertise, contextual sensitivity, and domain-specific knowledge.

Recommendations to improve translation quality in scientific domains:

- Incorporate **domain-specific translation memory** into CAT tools.
- Cross-check terms using validated **terminological databases** like IATE or Termium.
- Combine **machine translation** for initial drafts with **human post-editing** for accuracy (*Koehn, 2020*).
- Encourage **translator training** in specialized domains to ensure contextual awareness.

Computational tools significantly improve translation productivity. However, they are not substitutes for the nuanced understanding of human translators. Translation is not merely about word-for-word substitution but requires sensitivity to semantic, pragmatic, and cultural context. A hybrid model—technology-assisted, human-reviewed—emerges as the most effective strategy for translating scientific terminology.

References:

- Newmark, P. (1988). *A Textbook of Translation*. Prentice Hall. → cited in Sections 1, 2, 4, and 6
- Vinay, J.P. & Darbelnet, J. (1958). *Comparative Stylistics of French and English*. → cited in Section 2
- Vermeer, H.J. (1989). *Skopos and Commission in Translational Action*. → cited in Section 2

• Hutchins, J. (2005). *Computer-Assisted Translation: The State of the Art*. → cited in Section 3

• Bowker, L. & Fisher, D. (2010). *Computer-aided Translation Technology*. University of Ottawa Press. → cited in Sections 3 and 5

• Koehn, P. (2020). *Neural Machine Translation*. Cambridge University Press. → cited in Sections 4 and 7